Exhibit IND10

Exhibit I

[To N. Bryan's Opening Expert Report]

Effects of Noise, Image Resolution, and ROI Definition on the Accuracy of Standard Uptake Values: A Simulation Study

Ronald Boellaard, PhD; Nanda C. Krak, MD; Otto S. Hoekstra, PhD; and Adriaan A. Lammertsma, PhD

PET Center, VU University Medical Center, Amsterdam, The Netherlands

Semiquantitative standard uptake values (SUVs) are used for tumor diagnosis and response monitoring. However, the accuracy of the SUV and the accuracy of relative change during treatment are not well documented. Therefore, an experimental and simulation study was performed to determine the effects of noise, image resolution, and region-of-interest (ROI) definition on the accuracy of SUVs. **Methods:** Experiments and simulations are based on thorax phantoms with tumors of 10-, 15-, 20-, and 30-mm diameter and background ratios (TBRs) of 2, 4, and 8. For the simulation study, sinograms were generated by forward projection of the phantoms. For each phantom, 50 sinograms were generated at 3 noise levels. All sinograms were reconstructed using ordered-subset expectation maximization (OSEM) with 2 iterations and 16 subsets, with or without a 6-mm gaussian filter. For each tumor, the maximum pixel value and the average of a 50%, a 70%, and an adaptive isocontour threshold ROI were derived as well as with an ROI of 15 × 15 mm. The accuracy of SUVs was assessed using the average of 50 ROI values. Treatment response was simulated by varying the tumor size or the TBR. **Results:** For all situations, a strong correlation was found between maximum and isocontour-based ROI values resulting in similar dependencies on image resolution and noise of all studied SUV measures. A strong variation with tumor size of ≥50% was found for all SUV values. For non-smoothed data with high noise levels this variation was primarily due to noise, whereas for smoothed data with low noise levels partial-volume effects were most important. In general, SUVs showed under- and overestimations of ≥50% and depended on all parameters studied. However, SUV ratios, used for response monitoring, were only slightly dependent of ROI definition but were still affected by noise and resolution. **Conclusion:** The poor accuracy of the SUV under various conditions may hamper its use for diagnosis, especially in multicenter trials. SUV ratios used to measure response to treatment, however, are less dependent on noise, image resolution, and ROI definition. Therefore, the SUV might be more suitable for response-monitoring purposes.

Key Words: standard uptake value; response monitoring; region of interest; image resolution; noise

In nuclear medicine, PET with ^{18}F-FDG plays an important role. Several studies have shown its usefulness for staging and for measuring tumor treatment response ([1–8]). Semiquantitative analysis of ^{18}F-FDG uptake using standard uptake values (SUVs) may allow for a more precise diagnosis than visual assessment and the relative change in SUV has been used as a measure of treatment response ([5,6,9,10]).

The first step in determining the SUV is to derive the activity concentration (AC) in the tumor. Usually, the AC is obtained by placing a region of interest (ROI) over the tumor either visually ([11]), automatically using a threshold value ([12–16]), or using a fixed size ([10]). Unfortunately, various ROI methods are in use, making it difficult to compare different studies. Consequently, results obtained in one institution may not apply to results from other institutions. Other factors, such as reconstruction algorithm and filter, scanner sensitivity and scan duration, sinogram noise, and partial-volume effects, can also affect the accuracy of the measured AC ([17–25]).

The purpose of this study was to evaluate the effects of noise, image resolution, and ROI definition on the accuracy of measured SUVs. To this end, both simulation and experimental studies were performed, allowing us to separately investigate the effects of tumor size, tumor-to-background ratio (TBR), noise, image resolution, and ROI definitions on the accuracy of the measured AC. Furthermore, the effects on the accuracy of the observed relative changes in SUV were also investigated.

MATERIALS AND METHODS

First, a phantom study was performed to verify the results obtained with the simulation studies. Next, 2 simulation studies were performed. In the first study, a range of tumor characteristics and scanning or reconstruction conditions were simulated. The second study was performed to determine the relation between noise and bias of the maximum pixel value within a tumor.

Phantom Study

An anthropomorphic thorax phantom (Data Spectrum) containing 2 lungs and a liver insert was used. The large background compartment (soft tissue) was filled with an ^{18}F-FDG solution of 5 kBq/mL. Spheres, representing tumors of 8, 12, and 30 mm, were

filled with an ^{18}F-FDG solution of 20 kBq/mL (TBR = 4) and positioned in the mediastinum region of the phantom. Twenty-five 2-dimensional (2D) emission scans, each of 900 kilocounts, corresponding to the average number of counts observed in 5-min patient scans, were acquired using an ECAT HR+ scanner (CTI/ Siemens) (26,27). By performing acquisitions terminated on the number of acquired counts, statistically equivalent sinograms were obtained, allowing for reproducibility assessment. A 5-min transmission scan was acquired for attenuation correction. Data were reconstructed using ordered-subset expectation maximization (OSEM) with 2 iterations and 16 subsets (18). Gaussion filters of 0- and 6-mm full width at half maximum (FWHM) were used to match the image resolutions used in the simulations.

Simulation Study I: SUV Accuracy and Relative Change

A mathematic phantom was derived from a 2D dynamic ^{18}F-FDG PET scan of a patient. The patient data were summed from 15 to 45 min, and the resulting sinogram was reconstructed using OSEM with 2 iterations and 16 subsets and postsmoothed using a 6-mm FWHM gaussian filter. Subsequently, pixel values were scaled to an AC of ~5 kBq/mL, corresponding to the average soft-tissue AC in patient studies. In this mathematic phantom, tumors were simulated with spheres. Tumor sizes of 10-, 15-, 20-, and 30-mm diameter were used to cover the lower part of the clinically relevant range, as determination of the SUV is most challenging for small tumors. Tumors were located in the breast and lung. TBRs of 2, 4, and 8 (10, 20, and 40 kBq/mL) were applied. The size of the mathematic phantom corresponds to that of a patient of about 80 kg.

Next, noise-free sinograms were generated by forward projection of the image. These sinograms represented the number of true coincidences. Random and scattered coincidences were added to obtain prompts. Randoms were assumed to be distributed uniformly over the sinogram. Scattered coincidences were derived from forward projecting the difference between a scatter-corrected and a noncorrected image. Poisson noise was added to all sinograms. True coincidence sinograms were generated by subtracting the noisy random and scatter sinograms from the noisy prompts sinograms. Three noise levels were simulated corresponding to 2D data obtained for 3–5, 7–10, and 30–40 min or 0.75E+7, 1.5E+7, and 6.0E+7 noise equivalent counts (NEC) (28), respectively. For each combination of tumor size, TBR, and noise level, 50 noisy sinograms were generated to evaluate reproducibility and bias of the SUVs.

All sinograms were reconstructed using OSEM with 2 iterations and 16 subsets (ECAT version 7.2; CTI/Siemens). Image matrix sizes of 128 × 128 or 256 × 256, corresponding to pixel sizes of 5.12 × 5.12 and 2.56 × 2.56 mm, were used. Reconstructed images were postsmoothed using a gaussian filter such that image resolutions equaled 5- and 8-mm FWHM.

Finally, the phantom was adjusted by 25% to simulate patient weights of 60 and 100 kg, respectively. The number of acquired counts was adjusted using an empirically derived relation between patient weight and the counts-per-minute emission scan.

Simulation Study II: Noise and Bias

To assess the relations between SUV, tumor size, and noise, a simulation was performed using a uniform phantom containing an AC of 5 kBq/mL. Various noise levels, obtained by adding Poisson noise to the simulated sinograms, were applied, resulting in coefficients of variation (COVs) of ~0%, 10%, 20%, 35%, and 50%.

For each noise level, 100 simulations were performed. Spheric 3-dimensional (3D) ROIs of 11, 15, 21, 32, 39, and 50 mm in diameter were projected onto these images, and the maximum pixel value within each ROI was derived. The average maximum pixel value and its SE for each ROI size over the 100 simulations per noise level were calculated. These average values were normalized to true AC (5 kBq/mL).

Data Analysis

The same ROI methods were used for the phantom and the simulation study:

- 3D isocontour at 50% of maximum pixel value within tumor (ROI^{50});
- 3D isocontour at 70% of maximum pixel value (ROI^{70});
- 3D isocontour half way between background and maximum pixel value (ROI$^{0.5(Max+BG)}$);
- Maximum pixel value only;
- 15 × 15 mm square ROI centered on the location of maximum pixel value (ROI$^{15\times15}$).

These ROIs cover the various types of ROIs that are in regular use (12–15,29). Manual definition of ROIs (30) was not attempted, because it was impossible (~21,600 ROIs) for the present study. Moreover, other studies have shown that automatic definition of the ROI improves the interobserver reproducibility and accuracy of the measured AC (14). ROI$^{0.5(Max+BG)}$ was chosen because this ROI theoretically corresponds most closely to the actual tumor size. Furthermore, definition of ROI^{50} is not always possible for small tumors with low uptake, where ROI^{50} is lower than the background AC. ROI$^{0.5(Max+BG)}$ avoids this problem. ROI$^{15\times15}$ was included as a representative of a fixed-sized ROI (10).

Fifty AC values for the simulation study data and 25 for the phantom data were obtained for each combination of ROI, tumor size, TBR, tumor location, noise level, reconstruction matrix, and resolution. The accuracy of the observed AC was obtained from the average value over 50 simulated or 25 measured data. The recovery coefficient is defined here as the ratio between the observed AC and the true simulated AC. Note that the recovery coefficient may be larger than 1.0, indicating an overestimation of the actual AC.

The accuracy of measured relative changes, simulating treatment response, was investigated using the same simulation. Two situations were considered: (a) tumors of 10- and 30-mm diameter with a decrease in uptake, but no variation in size; and (b) tumors with TBR = 2 or 8 with a decrease in tumor size, but not in ^{18}F-FDG uptake.

Definition of "Defaults"

An extensive amount of data was generated by the simulations and phantom studies. Therefore, unless stated otherwise, results presented here are limited to a default situation:

- TBR = 4;
- Patient weight of 80 kg;
- Noise level corresponding to 5- to 7-min 2D scans or 1.5E+7 NEC;
- OSEM reconstructions with matrix size = 128 × 128 and 6-mm FWHM gaussian smoothing.

RESULTS

Phantom Study

Figure 1A shows the recovery coefficient as a function of tumor size using different ROIs. In Figure 1B, corresponding data are shown for the simulation study. Recovery coefficients increase with tumor size for all ROI methods. Use of the maximum pixel value did not reduce the variation of measured AC with tumor size. Recovery coefficients obtained with $ROI^{0.5(Max+BG)}$ were almost equal to ROI^{50} data for large tumors and to ROI^{70} data for small tumors.

In Figures 1C and 1D, the AC measured with threshold and 15×15 mm ROIs is plotted against the maximum pixel value for each of the 25 measurements and 50 simulations for all tumor sizes. A strong correlation among various measures of AC can be seen. Second, reproducibilities are similar for the maximum pixel value and all mean values based on threshold-based isocontour ROIs (data are almost located on a straight line). Note that use $ROI^{15\times15}$ resulted in the largest underestimations of AC due to the relatively small tumor sizes applied in this study. Furthermore, $ROI^{15\times15}$ showed the poorest correlation with maximum pixel value. For small spheres, this ROI included too many background pixels and, in the remainder of this study, it will only be used for the assessment of the accuracy of relative change. In general, consistent variations of SUV with the sphere size and ROI method were found between phantom and simulation studies, and the rest of this article focuses on simulation data only.

Simulation Study I: SUV Accuracy

Results obtained for breast tumors were almost identical to lung tumor data. For lung tumors, recovery coefficients were slightly lower due to a larger contrast with background activity. Therefore, results are presented for simulated breast tumors only.

Figures 2A–2D show recovery coefficients for data obtained with various noise levels and different image resolutions. For clarity, data based on maximum pixel value and ROI^{50} are presented, as intermediate results were obtained for other thresholds. Figure 2A shows that use of maximum pixel values resulted in large overestimations of the AC, which increased with higher noise levels. A similar variation with noise levels was found for ROI^{50} values (Fig. 2B), but these were more in agreement with the actual AC. For smoothed data (Figs. 2C and 2D), differences in recovery coefficients between various noise levels were much smaller, due to the noise-suppressing effect of smoothing. Substantial variations of recovery coefficient with tumor size were found for both smoothed and non-smoothed data.

Figures 3A–3D demonstrate the effect of isocontour values on the recovery coefficient and its variation with tumor size for TBR equal to 2 and 8 and for 2 different image resolutions. For TBR = 2, ROI^{50M-BG} agreed most with ROI^{70} data; for TBR = 8, the best agreement with ROI^{50} was found. Furthermore, for smoothed data, ROI^{50} could not be defined for the smallest 2 tumors of 10 and 15 mm in the case of TBR = 2, because it included the entire phantom.

Image matrix (pixel size) affects both noise and image resolution. Data presented in Table 1 indicate that the AC obtained with maximum pixel values and ROI^{50} strongly depends on image matrix size and width of the smoothing kernel. For data with an image resolution of 8-mm FWHM,
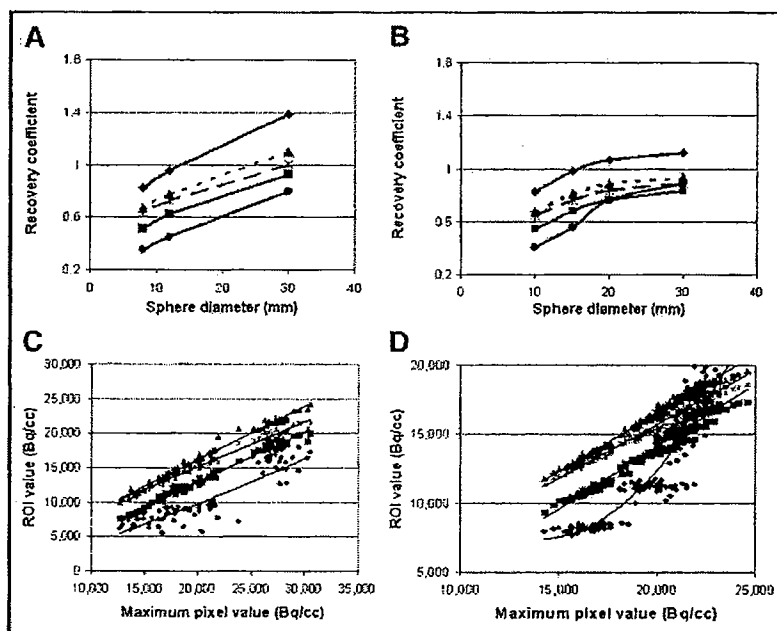

FIGURE 1. Recovery coefficients as function of sphere size for phantom data (A) and simulation data (B) and ROI values as function of maximum pixel value for phantom data (C) and simulated data (D). ◆ (A and B only) = maximum pixel value; ▲ = ROI^{70}; ✶ = $ROI^{0.5(BG+MAX)}$; ■ = ROI^{50}; ● = $ROI^{15\times15}$.

FIGURE 2. Recovery coefficients as function of sphere size for various noise levels. ◆ = low NEC; ▲ = medium NEC; ■ = high NEC. (A and B) No smoothing of images. (C and D) Smoothing of images. A and C represent maximum pixel value; B and D represent ROI50. Error bars equal 1 SD.
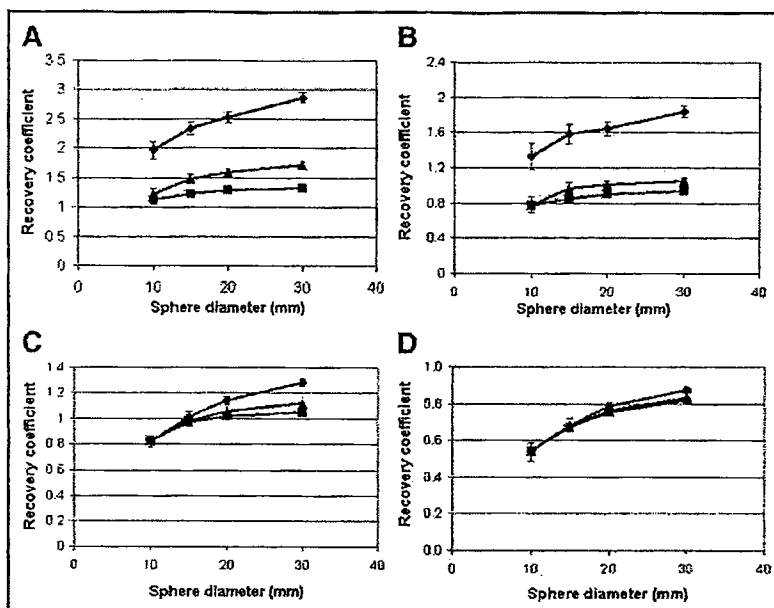
the effect of pixel size is smaller than for those with a resolution of 5-mm FWHM.

The effects of patient weight are given in Table 2. Up to ~20% higher and ~10% lower ACs were found at a 25% decrease or increase of patient weight. For smoothed data, differences are <5%.

Simulation I: Relative Change or Response Monitoring

Relative changes of SUV as a function of the type of ROI are presented in Figure 4A for a 10-mm-diameter tumor and in

Figure 4B for a 30-mm-diameter tumor, both for 5-mm FWHM image resolution data. Note that responses of 50% and 25% are defined here as residual uptakes—that is, as 50% and 75% decreases in ^{18}F-FDG uptake compared with the initial concentration of 40 kBq/mL. Figures 4C and 4D show the same data obtained with an image resolution of 8-mm FWHM. Most measured responses were similar and independent of image resolution. However, in the case of 8-mm FWHM image resolution, a response could not be obtained

FIGURE 3. Recovery coefficients for various ROIs. ◆ = maximum pixel value; ▲ = ROI70; ● = ROI0.5(BG+MAX); ■ = ROI50. A and C represent data for TBR = 2; B and D represent data for TBR = 8. (A and B) Nonsmoothed data. (C and D) Smoothed data. SEs varied from ~9% to ~4% and from ~5% to ~2% for 10- to 30-mm spheres for all ROI methods for nonsmoothed (A and B) and smoothed data (C and D), respectively.
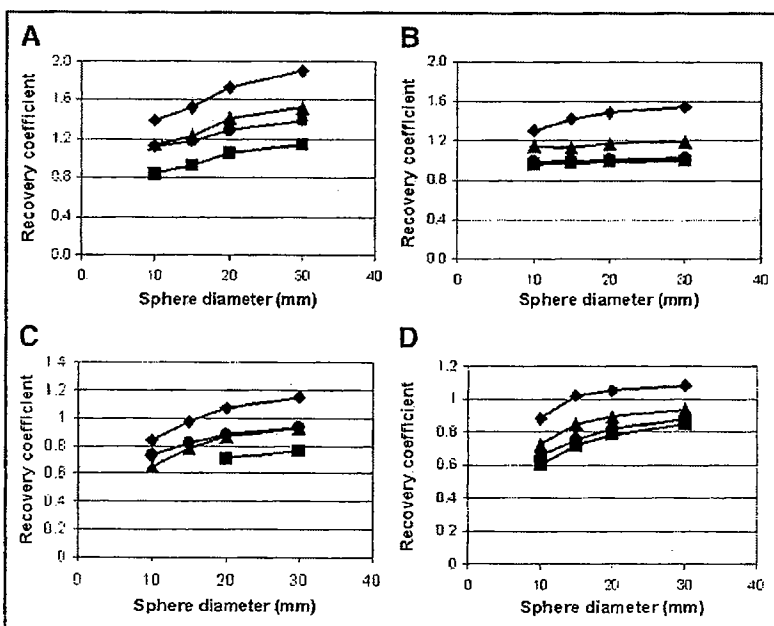
TABLE 1
Effects of Image Resolution and Matrix Size on Recovery Coefficient as Function of Sphere Size

	Maximum pixel value				ROI[50]			
Image resolution								
FWHM (mm)	5	5	8	8	5	5	8	8
Image matrix size	128	256	128	256	128	256	128	256
Sphere size								
10	1.05 (0.14)	1.30 (0.15)	0.77 (0.03)	0.82 (0.05)	0.71 (0.14)	0.86 (0.16)	0.51 (0.03)	0.54 (0.05)
15	1.17 (0.10)	1.48 (0.11)	0.94 (0.03)	0.99 (0.04)	0.81 (0.08)	0.96 (0.10)	0.65 (0.02)	0.68 (0.04)
20	1.25 (0.07)	1.59 (0.07)	1.03 (0.02)	1.06 (0.03)	0.88 (0.05)	1.01 (0.05)	0.73 (0.02)	0.76 (0.02)
30	1.37 (0.05)	1.72 (0.07)	1.09 (0.02)	1.12 (0.03)	0.93 (0.03)	1.05 (0.04)	0.81 (0.01)	0.83 (0.01)

SEs are given in parentheses.

using ROI[50] for a SUV decrease to 25%, because this ROI included the entire patient. ROI[15×15] provided the poorest estimate of SUV response for the smaller tumor sizes. Error bars in Figures 4 illustrate the effects of ROI type on the reproducibility (1 SD) of the measured response. Similar reproducibilities among most ROI types were observed, except for ROI[15×15].

When only tumor size changed, an artifactual SUV response could be expected due to partial-volume effects. The latter is shown in Figures 5A–D. For all ROIs, similar erroneous responses were measured, which depended on the TBR, tumor size, and image resolution. Errors bars have been omitted for clarity, but reproducibilities show a dependence on ROI type similar to that presented in Figure 4.

Figures 6A–6D show the effect of noise on the measured responses using maximum pixel values. Similar results were obtained for all other ROIs.

Simulation II: Noise and Bias

In Figure 7A the recovery coefficient, defined here as the ratio of the average maximum pixel value over 100 simulations and the uniform AC value, is presented as a function of the ROI size for various noise levels. Increasing bias with

the noise level and the ROI size was found. This indicates an increased probability to observe higher maxima with an increasing amount of pixels. In Figure 7B, the data for a background noise level of 35% are shown. Errors bars illustrate the variation of reproducibility (1 SD) with ROI size.

DISCUSSION

Applicability and Limitations of Simulations

Using simulation and phantom studies it is possible to obtain data under well-defined and clinically relevant conditions. However, some limitations apply: (a) tumors were represented using 3D spheres, which is a simplification; (b) the AC within a tumor was constant, thereby excluding effects of tumor heterogeneity. It can be argued, however, that results obtained for homogeneous spheric tumors are "ideal" and that heterogeneous nonspheric tumors will result in additional "noise" on top of the present results. However, clinical evaluation of ROI and reconstruction methods is still required, which is part of ongoing research.

In addition, the simulation focused on tumors located in

TABLE 2
Effects of Image Resolution and Patient Weight on Recovery Coefficient as Function of Sphere Size

|	Maximum pixel value			Maximum pixel value			ROI[50]			ROI[50]		
Image resolution												
FWHM (mm)	5	5	5	8	8	8	5	5	5	8	8	8
Patient weight (kg)	60	80	100	60	80	100	60	80	100	60	80	100
Sphere size												
10	0.96	1.05	1.20	0.75	0.77	0.77	0.74	0.86	1.00	0.50	0.51	0.50
	(0.17)	(0.16)	(0.17)	(0.06)	(0.08)	(0.09)	(0.19)	(0.17)	(0.27)	(0.05)	(0.08)	(0.09)
15	1.10	1.17	1.43	0.94	0.94	0.96	0.89	0.96	1.24	0.65	0.65	0.65
	(0.14)	(0.16)	(0.15)	(0.07)	(0.06)	(0.08)	(0.13)	(0.16)	(0.18)	(0.05)	(0.05)	(0.07)
20	1.12	1.25	1.45	1.01	1.03	1.06	0.88	1.01	1.30	0.73	0.73	0.75
	(0.12)	(0.14)	(0.13)	(0.04)	(0.05)	(0.05)	(0.14)	(0.14)	(0.14)	(0.03)	(0.03)	(0.06)
30	1.25	1.37	1.64	1.05	1.09	1.13	0.93	1.05	1.40	0.80	0.81	0.82
	(0.11)	(0.11)	(0.11)	(0.03)	(0.04)	(0.03)	(0.12)	(0.09)	(0.12)	(0.01)	(0.02)	(0.02)

SEs are given in parentheses.

FIGURE 4. Measured residual uptake using various ROIs for simulated tumors showing a decrease in uptake only. (A and B) Nonsmoothed data. (C) Smoothed data. A and C give response of tumors with constant tumor size of 10-mm diameter; B and D show response of tumors with constant tumor size of 30 mm. Residual uptake was measured using maximum pixel value, ROI^{50}, ROI^{70}, $ROI^{0.5(BG+MAX)}$, and $ROI^{15 \times 15}$ as shown from left to right in each set of 5 bars. Error bars represent 1 SD. Missing bars in this figure correspond to missing data due to impossible ROI definition—that is, ROI enclosed entire patient.
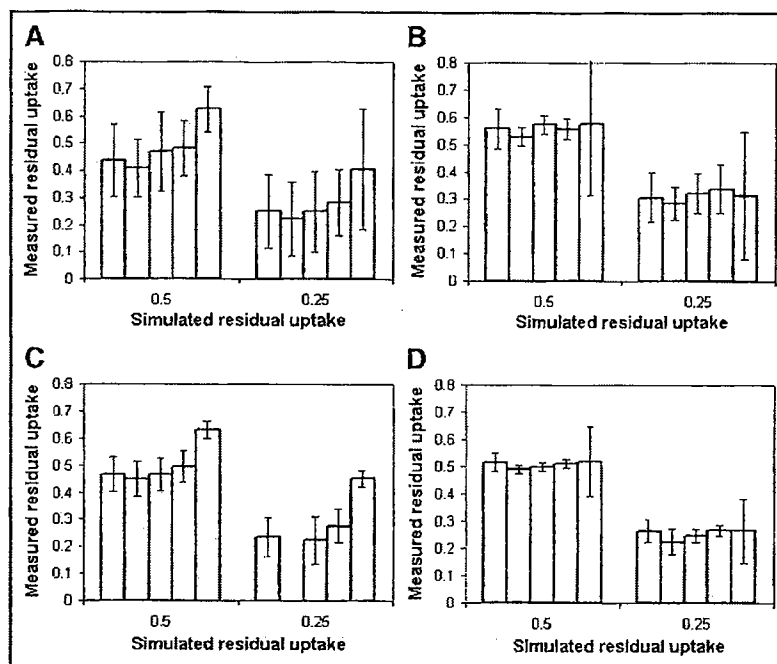
the thorax. When studies are performed at other body regions, somewhat different recovery coefficients can be expected due to differences in noise level resulting from differences in attenuation losses, random and scatter contribution. However, the variation of the SUV with resolution, noise, and ROI method will be similar, because noise, resolution, and ROI method are the underlying causes of the observed variations.

FIGURE 5. Measured residual uptake using various ROIs keeping tumor uptake constant. Max = maximum pixel value; 50% = ROI^{50}; 70% = ROI^{70}; 0.5BGM = $ROI^{0.5(BG+MAX)}$; 15×15 = $ROI^{15 \times 15}$. (A and B) No smoothing of image. (C and D) Smoothing of image. A and C represent tumor data with TBR = 2; B and D represent tumor data with TBR = 8. Reduction of diameter to 10, 15, and 20 mm and not at all (it stays 30 mm) is shown from left to right in each set of 4 bars. Missing bars in this figure correspond to missing data due to impossible ROI definition—that is, ROI enclosed entire patient.
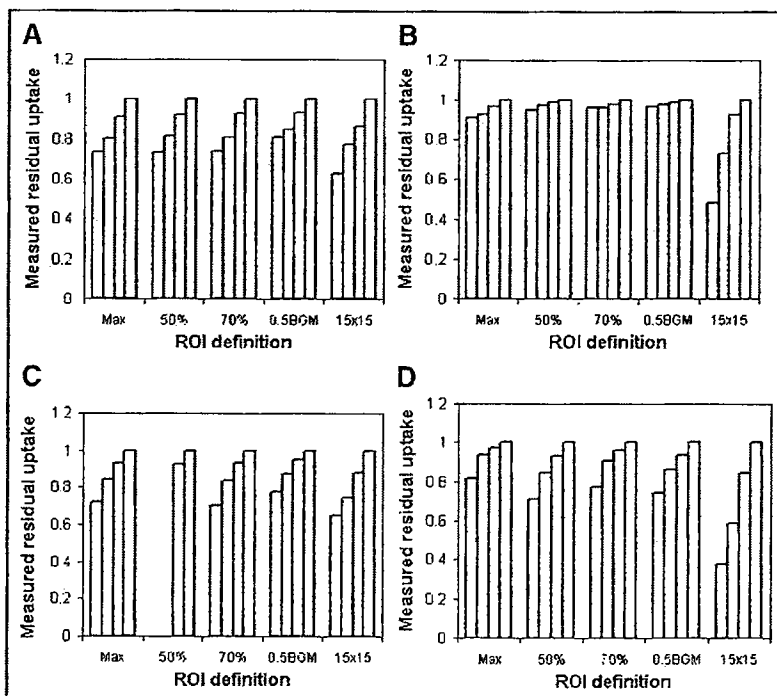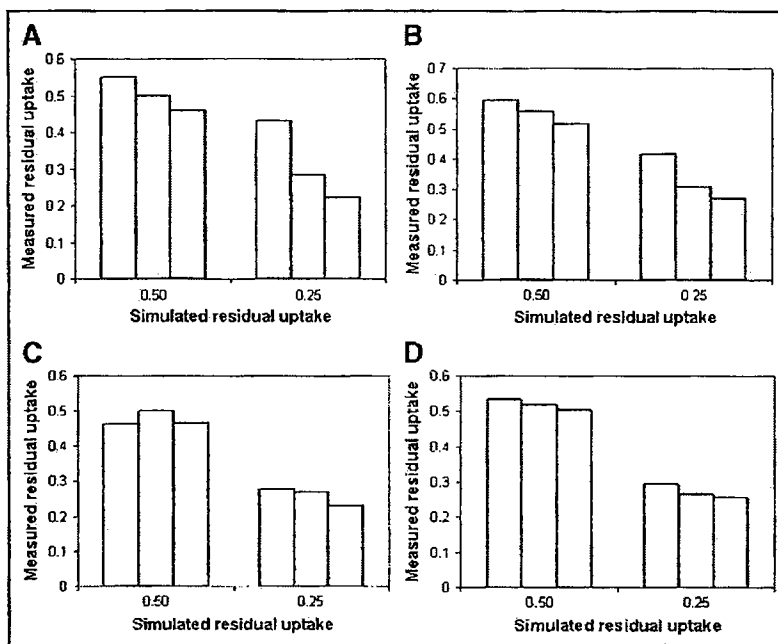
FIGURE 6. Effects of sinogram statistics on measured residual uptake for non-smoothed data (A and B) and smoothed data (C and D), assuming constant tumor size of 10 mm (A and C) and 30 mm (B and D). Low, medium, and high NEC data are shown from left to right in each set of 3 bars.

Accuracy of SUV

The observed recovery coefficients showed similar variation with sphere size under various conditions for all ROIs evaluated. There is, however, a difference for smoothed and unsmoothed data. In Table 3, values for the COV of background pixels are given for both smoothed and nonsmoothed data for each noise level. It can be deduced that image noise varies more with sinogram statistics for nonsmoothed data than for smoothed data. The relationship between image noise and sinogram statistics is consistent with those observed by Boellaard et al. (*18*). The results presented in Figure 2 indicate that the maximum pixel value increases both with ROI or tumor size and with noise level. Equivalently, maximum pixel values within an object increase with image noise level and with increasing object size. Similar results were observed by Falen et al. (*21*), who found SUV increases up to 70% with a higher number of iterations. Increasing the number of iterations not only improves convergence but also increases image noise. There-

fore, images with higher noise show more positive bias for both the maximum pixel and the ROI value within a sphere (Fig. 7), and this bias increases with sphere size. This explains the large variation of maximum pixel value with object size for nonsmoothed data. Note that the variation of the maximum pixel value for nonsmoothed data is similar to that for smoothed data, indicating that partial-volume effects have only a minor contribution to the observed dependence on sphere size. Smoothing the data reduces effects of image noise but introduces a larger partial-volume effect. Consequently, variations of the maximum pixel value with sphere or tumor size are substantial for both smoothed and nonsmoothed images, but the underlying mechanisms are different. Because all isocontour-based ROI methods use the maximum pixel value as a reference, similar results were found using these ROIs (Figs. 1C and 1D).

The effects of the ROI method on the accuracy of recovery coefficients are trivial. ROI[70] and ROI[50] values are about
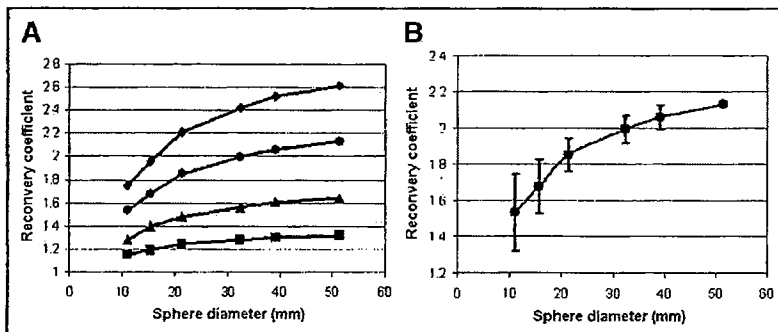

FIGURE 7. (A) Recovery coefficients determined using maximum pixel value within ROI as function of ROI size for various noise levels. ■ = COV of 10%; ▲ = COV of 20%; ● = COV of 35%; ◆ = COV of 50%. (B) Recovery coefficients based on maximum pixel value as function of ROI size for noise level = 35%. Error bars now illustrate variation of reproducibility (1 SD) as function of ROI size.

TABLE 3
COV of Pixel Values in Homogeneous Background of
Simulated Phantom as Function of Sinogram Statistics
and Image Resolution

Duration (min)	NEC	COV (%)	
		5-mm resolution	8-mm resolution
3–5	0.75E+7	50	16
7–10	1.5E+7	33	12
30–40	6.0E+7	17	6

15% and 30% lower than the maximum pixel value. $ROI^{0.5(Max+BG)}$ values show intermediate results, but its value was in general close to the ROI^{50} values for tumors with a high TBR and closer to ROI^{70} for a low TBR. Smoothing the data led to an overall reduction of all measures. ROI^{50} seemed to be most accurate for high image resolution (noisy) data, whereas the maximum pixel value was more accurate for smoothed (low noise) data. $ROI^{15×15}$ data provided the worst results among all ROIs tested, because the data included too many background pixels. It is therefore not a good candidate ROI for determining the average AC in small tumors (<20 mm).

The Nyquist principle requires the pixel size to be at least 2 times smaller than image resolution. For the simulated images with a resolution of 5-mm FWHM, a matrix size of 128 × 128 or a pixel size of 5 × 5 × 2.5 mm does not fulfill this criterion and will reduce image resolution (Table 1). Reducing the pixel size to 2.5 × 2.5 mm does no longer violate the Nyquist principle and will therefore not degrade image resolution. As expected, for smoothed data, with an image resolution of 8-mm FWHM, the effect of matrix or pixel size was much smaller.

Finally, patient weight has an effect on the accuracy of the measured AC. Variation in weight causes variation in sinogram counts or noise (Table 2), which causes bias. When data are sufficiently smoothed, the effects of noise are minimized at the cost of an increased partial-volume effect. A useful strategy to overcome this problem might be to adjust the dose to patient weight.

Response Monitoring

On average, the measured response was almost independent of the ROI method when the tumor size was kept constant. The good agreement of the measured response among all ROIs can be understood from the data presented in Figures 1C and 1D, which illustrate the strong correlation of ROI values over all noisy realizations. Therefore, the reproducibility and accuracy of the AC of the maximum pixel value in fact determines the reproducibility and accuracy of the AC measured with an isocontour ROI. Consequently, similar responses were measured with all isocontour-based ROIs. A slightly lower accuracy and very poor reproducibility were observed for $ROI^{15×15}$ data because this region is too large for the small tumor sizes investigated in this study.

The measured response for tumors that showed a large variation in size strongly depended on tumor size, which is a logical consequence of the data presented in Figures 1A and 1B and in Figures 7A and 7B. Smoothing the data reduced, but did not remove, the differences among the ROIs. As mentioned previously, both noise-induced bias for nonsmoothed data and partial-volume effects for smoothed data are the underlying causes of the observed effects.

The (almost) independence of the measured response on noise, resolution, and ROI method is explained by the fact that most factors cancel out when calculating SUV ratios, such as bias due to noise; absolute quantification due to the applied ROI method; patient weight (= noise); image resolution. When the tumor size does not vary much, partial-volume effects will also cancel out (at least partially). However, more accurate results can be expected when accurate partial-volume corrections can be applied. Note, however, that smoothing improves the reproducibility of the observed responses.

Considerations for Use of SUV in Clinical Practice

From the present study it can be concluded that the SUV obtained under specific conditions may not be compared directly with those obtained, reconstructed, or analyzed under other conditions (*31*). Consistency of data acquisition and analysis protocols is therefore required. This can easily be adhered to within one institution, explaining the successful use of the SUV in differentiating benign from malignant tumors. However, SUV measures will probably vary strongly among institutions. Therefore, SUV threshold values used to differentiate between malignant and benign lesions should not be taken from the literature and used without validating the appropriateness. Moreover, another important observation is made by Thie et al. (*32*), who found that SUVs show a lognormal rather than a gaussian distribution, which will require a new review of applied SUV thresholds. Finally, standardization of imaging and analysis protocols is required for multicenter studies. Standardization is required, for example, for acquisition mode; average NEC (combination of scanner sensitivity and scan durations); reconstruction method and image resolution; interval between ^{18}F-FDG administration and scanning; partial-volume corrections; head or feet at first bed position; ROI method; emptying of bladder; SUV calculation method. Any remaining differences may be determined and corrected for by including a standardized phantom experiment—for example, using anthropomorphic phantoms (Data Spectrum).

Another application of the SUV is to measure the treatment response of a tumor (*5,6,8,10*). Assuming consistent data processing for multiple scans of 1 patient, the measured response might be less dependent on noise, image resolution, and ROI definition (Figs. 4–6). Minimizing noise levels by smoothing reduces bias and seems to have a minimal effect on the accuracy of the measured response without the need for partial-volume correction algorithms.

Some partial-volume effects were observed, however, when the tumor volume changed drastically. This can only be solved by the introduction of partial-volume correction methods. These variations in tumor size, however, were most likely an extreme situation in relation to changes seen in clinical practice. Because of the relative independence of the SUV response with respect to several study parameters, multicenter response-monitoring studies should be feasible.

CONCLUSION

A phantom and simulation study was performed to determine the effect of noise, image resolution, and ROI definition on the accuracy of the SUV. The SUV depended strongly on all studied parameters, and it can only be used for diagnostic purposes when data acquisition and processing are performed in a standardized way. This might be a problem for multicenter studies. The SUV ratios used to measure the treatment response depended less on noise and image resolution and are therefore more suitable for multicenter trials.

ACKNOWLEDGMENTS

REFERENCES

1. Aoki J, Watanabe H, Shinozaki T, et al. FDG PET of primary benign and malignant bone tumors: standardized uptake value in 52 lesions. Radiology. 2001;219:774–777.
2. Aoki J, Watanabe H, Shinozaki T, et al. FDG-PET for preoperative differential diagnosis between benign and malignant soft tissue masses. Skeletal Radiol. 2003;32:133–138.
3. Avril N, Bense S, Dose J, et al. PET imaging of the breast: quantitative image analysis [abstract]. J Nucl Med. 1996;37(suppl):99P.
4. Dehdashti F, Siegel BA, Griffeth LK, et al. Benign versus malignant intraosseous lesions: discrimination by means of PET with 2-[F-18]fluoro-2-deoxy-D-glucose. Radiology. 1996;200:243–247.
5. Hoekstra CJ, Paglianiti I, Hoekstra OS, et al. Monitoring response to therapy in cancer using [F-18]-2-fluoro-2-deoxy-D-glucose and positron emission tomography: an overview of different analytical methods. Eur J Nucl Med. 2000;27:731–743.
6. Hoekstra CJ, Hoekstra OS, Stroobants SG, et al. Methods to monitor response to chemotherapy in non-small cell lung cancer with ¹⁸F-FDG PET. J Nucl Med. 2002;43:1304–1309.
7. Hubner KF, Buonocore E, Gould HR, et al. Differentiating benign from malignant lung lesions using "quantitative" parameters of FDG PET images. Clin Nucl Med. 1996;21:941–949.
8. Krak NC, van der Hoeven JJM, Hoekstra OS, Twisk JWR, van der Wall E, Lammertsma AA. Measuring [F-18]FDG uptake in breast cancer during chemotherapy: comparison of analytical methods. Eur J Nucl Med Mol Imaging. 2003;30:674–681.
9. Jansson T, Westlin JE, Ahlstrom H, Lilja A, Langstrom B, Bergh J. Positron emission tomography studies in patients with locally advanced and/or metastatic breast-cancer: a method for early therapy evaluation. J Clin Oncol. 1995;13:1470–1477.
10. Wahl RL, Zasadny K, Helvie M, Hutchins GD, Weber B, Cody R. Metabolic monitoring of breast-cancer chemohormonotherapy using positron emission tomography: initial evaluation. J Clin Oncol. 1993;11:2101–2111.
11. Avril N, Bense S, Ziegler SI, et al. Breast imaging with fluorine-18-FDG PET: quantitative image analysis. J Nucl Med. 1997;38:1186–1191.
12. Graham MM, Peterson LM, Hayward RM. Comparison of simplified quantitative analyses of FDG uptake. Nucl Med Biol. 2000;27:647–655.
13. Kole AC, Nieweg OE, Pruim J, et al. Standardized uptake value and quantification of metabolism for breast cancer imaging with FDG and L-[1-C-11]tyrosine PET. J Nucl Med. 1997;38:692–696.
14. Lee JR, Madsen MT, Bushnel D, Menda Y. A threshold method to improve standardized uptake value reproducibility. Nucl Med Commun. 2000;21:685–690.
15. Madsen MT, Lee JR, Bushnell DL, Menda Y. Standard uptake values using a fixed region threshold [abstract]. J Nucl Med. 2000;41(suppl):182P.
16. Weber WA, Ziegler SI, Thodunann R, Hanauske AR, Schwaiger M. Reproducibility of metabolic measurements in malignant tumors using FDG PET. J Nucl Med. 1999;40:1771–1777.
17. Akhurst T, Erdi Y, MacApinlac H, Humm J, Yeung H, Larson SM. A comparison of reconstruction algorithms with FDG-PET: iterative reconstruction segmented attenuation correction (IRSAC) allows greater lesion detectability than filtered back projection (FBP) [abstract]. Eur J Nucl Med. 1999;26:1161.
18. Boellaard R, van Lingen A, Lammertsma AA. Experimental and clinical evaluation of iterative reconstruction (OSEM) in dynamic PET: quantitative characteristics and effects on kinetic modeling. J Nucl Med. 2001;42:808–817.
19. Bruckbauer T, Casey M, Valk PE, Rao J, Finley BR, Farboud B. Optimizing 3D whole body acquisition for oncologic imaging on the ECAT ACCEL LSO PET system [abstract]. Eur J Nucl Med. 2001;28:1065.
20. Etchebehere ECSC, Macapinlac HA, Gonen M, et al. Qualitative and quantitative comparison between images obtained with filtered back projection and iterative reconstruction in prostate cancer lesions on F-18-FDG PET. Quart J Nucl Med. 2002;46:122–130.
21. Falen SW, Ivanovic M, Venkataraman S. Standard uptake value analysis: evaluation of acquisition and processing parameters for PET [abstract]. Mol Imaging Biol. 2003;5:106.
22. Hickeson M, Yun MJ, Matthies A, et al. Use of a corrected standardized uptake value based on the lesion size on CT permits accurate characterization of lung nodules on FDG-PET. Eur J Nucl Med Mol Imaging. 2002;29:1639–1647.
23. Lartizien C, Kinahan PE, Swensson R, et al. Evaluating image reconstruction methods for tumor detection in 3-dimensional whole-body PET oncology imaging. J Nucl Med. 2003;44:276–290.
24. Lonneux M, Borbath I, Bol A, et al. Attenuation correction in whole-body FDG oncological studies: the role of statistical reconstruction. Eur J Nucl Med. 1999;26:591–598.
25. Paul AK, Tatsumi M, Yutani K, Fujino K, Hashikawa K, Nishimura T. Effects of iterative reconstruction on image contrast and lesion detection in gamma camera coincidence imaging in lung and breast cancers. Nucl Med Commun. 2002;23:103–110.
26. Adam LE, Zaers J, Ostertag H, Trojan H, Bellemann ME, Brix G. Performance evaluation of the whole-body PET scanner ECAT EXACT HR+ following the IEC standard. IEEE Trans Nucl Sci. 1997;44:1172–1179.
27. Brix G, Zaers J, Adam LE, et al. Performance evaluation of a whole-body PET scanner using the NEMA protocol. J Nucl Med. 1997;38:1614–1623.
28. Strother SC, Casey ME, Hoffman EJ. Measuring PET scanner sensitivity: relating count rates to image signal-to-noise ratios using noise equivalent counts. IEEE Trans Nucl Sci. 1990;37:783–788.
29. Minn H, Zasadny KR, Quint LE, Wahl RL. Lung cancer: reproducibility of quantitative measurements for evaluating 2-[F-18]-fluoro-2-deoxy-D-glucose uptake at PET. Radiology. 1995;196:167–173.
30. Avril N, Rose CA, Schelling M, et al. Breast imaging with positron emission tomography and fluorine-18 fluorodeoxyglucose: use and limitations. J Clin Oncol. 2000;18:3495–3502.
31. Keyes JW Jr. SUV: standard uptake or silly useless value? J Nucl Med. 1995;36:1836–1839.
32. Thie JA, Hubner KF, Smith GT. The diagnostic utility of the lognormal behavior of PET standardized uptake values in tumors. J Nucl Med. 2000;41:1664–1672.